Grid continuation in the Entorhinal Cortex by multi-modal integration

Gábor Szirtes and András Lőrincz

Department of Information Systems, Eötvös Loránd University, Budapest, Hungary

szirtes@inf.elte.hu

1. Introduction

The hippocampal formation (HF; entorhinal cortex (EC), CA3,CA1,DG and the subiculum) plays a central role in *seemingly* distant memory functions (e. g. spatial and episodic memory) for which a coherent explanation has not yet been found. Some of the puzzling issues are 1, the relation between spatial and other forms of memories; 2, the role of the feedback loops in decoding-encoding; and 3, the nature of representations formed first in the HF then encoded into the neocortex (*what* and *how*).
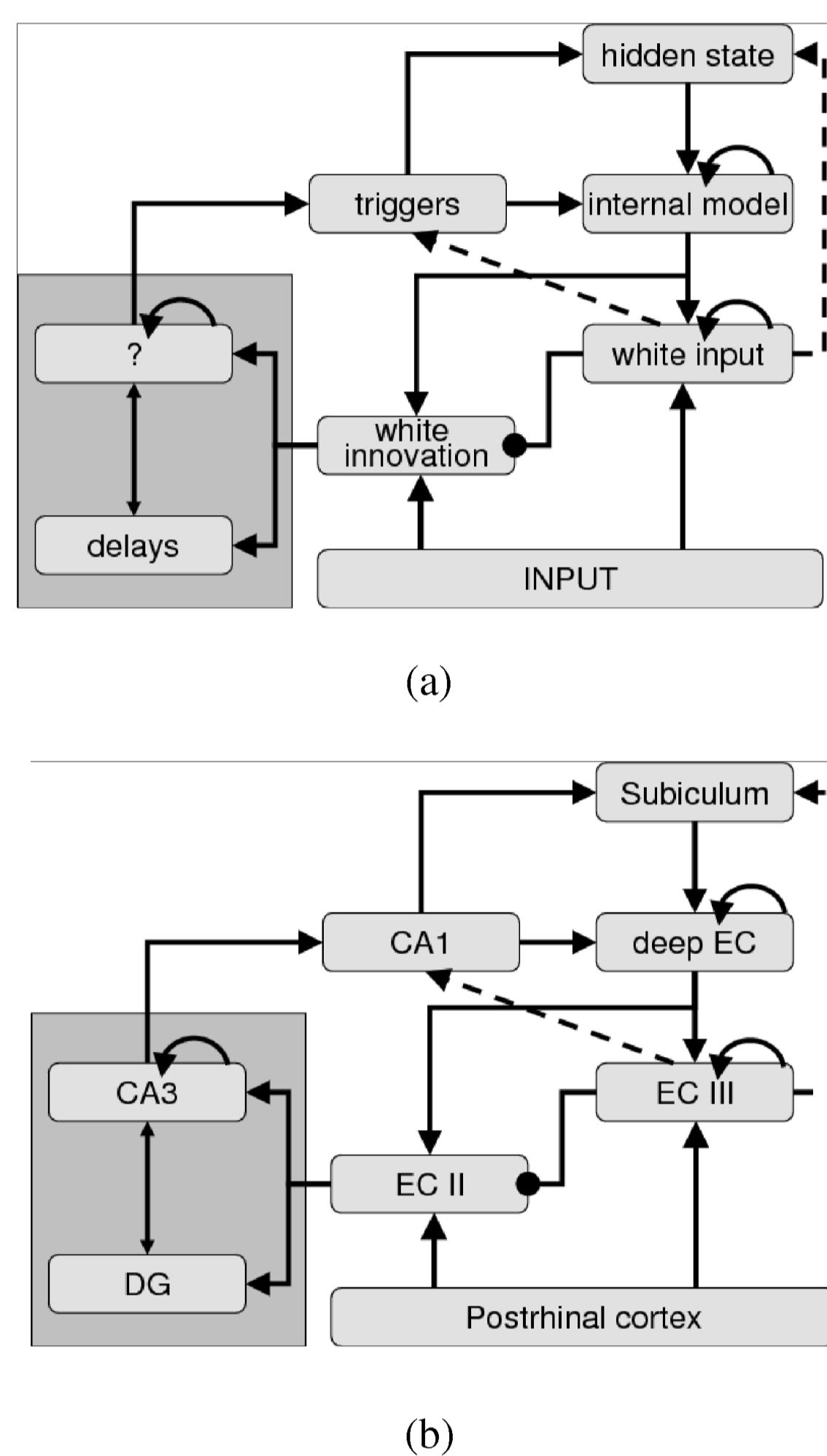

(a)

(b)

Figure 1: *(a) The connectionist architecture can identify a hidden ARMA(p,q) dynamical system driven by* **independent** *sources. The overall structure is a* **consequence** *of the suggested algorithm and the constraint of* **Hebbian** *(local) interactions. (b) Gross anatomy of HF. Dashed lined denote connections tuned in a supervised manner. Arrowhead lines denote mostly excitatory, circle head denotes inhibitory connections.*

We have recently suggested [1] that the core function of

HF is *system identification*: it is tuned to expose the hidden (and presumably) low dimensional dynamical system behind the observations. The emerging *internal* model can be used to predict and correct subsequent (noisy or partial) observations thus decreasing the computational load. In particular, the overall goal is to learn causal relationships in which **independent** driving sources can trigger events defined as change in the state. For **hidden** ARMA(p,q) models driven by independent sources, it can be proved [2] that model parameters (of the underlying dynamical system and the observation matrix) as well as the hidden sources can be recovered (up to some ambiguity). The following computational steps are needed: 1, estimating **innovation**, 2 **whitening** (decorrelation and normalization), **source separation** (e.g. by Independent Component Analysis, [3]) and **supervised learning** to form the internal model. Imposing locality on the computations (Hebbian constraint) resulted in a connectionist architecture (Fig 1.a) bearing strong similarities with the gross anatomy of HF Fig. 1.b.

When the generic model was applied on signals with implicit spatial information (random mixture of position dependent Gaussian blobs), it manifested grid- and place cell like spatial activity distribution in modules corresponding to the entorhinal layers and CA1, respectively. For mixtures of purely position **and** purely direction dependent inputs the model discovered the two independent subspaces of position and direction. For **conjunctive** inputs, separation discretized the 3D space of space and direction and ICA on novelty decreased the direction dependence (results not shown).

2. Grid continuation by prediction

Regarding spatial memories, grid cells ([4]) of the EC are supposed to provide a metric required for reliable spatial navigation (e.g. in path integration, see [5]). However, 1, they can be transiently distorted by smooth changes in the environment [6] 2, metric can be based on other representations [7] and 3, navigation may not always require HF [8]. In addition, the generating mechanism is not known, there are several competing theories (e.g. attractor dynamics models, assuming strong recurrent connections [5] and

oscillatory models [9, 10]). The core model is based on statistical learning, behavioral aspects are not yet included. Grid continuation (maintaining grid representation while not all modalities can be accessed), for example, cannot be explained by pure statistics. The transient distortion of grids may imply that it cannot purely be generated by internal (sensory independent) mechanism. Instead, we suggest that a grid basis can be learnt based on the sensory statistics and then an internal model of the dynamics of the grid activity patterns is updated by sensory **and** motor signals. In this way -depending on the context- different modalities can be taken into account according to the extent the system trusts in their reliability (probabilistic switching). In accord with [11] we emphasize that integration should be realized by a "substitution" mechanism, in which modalities can replace each other in updating the internal model.
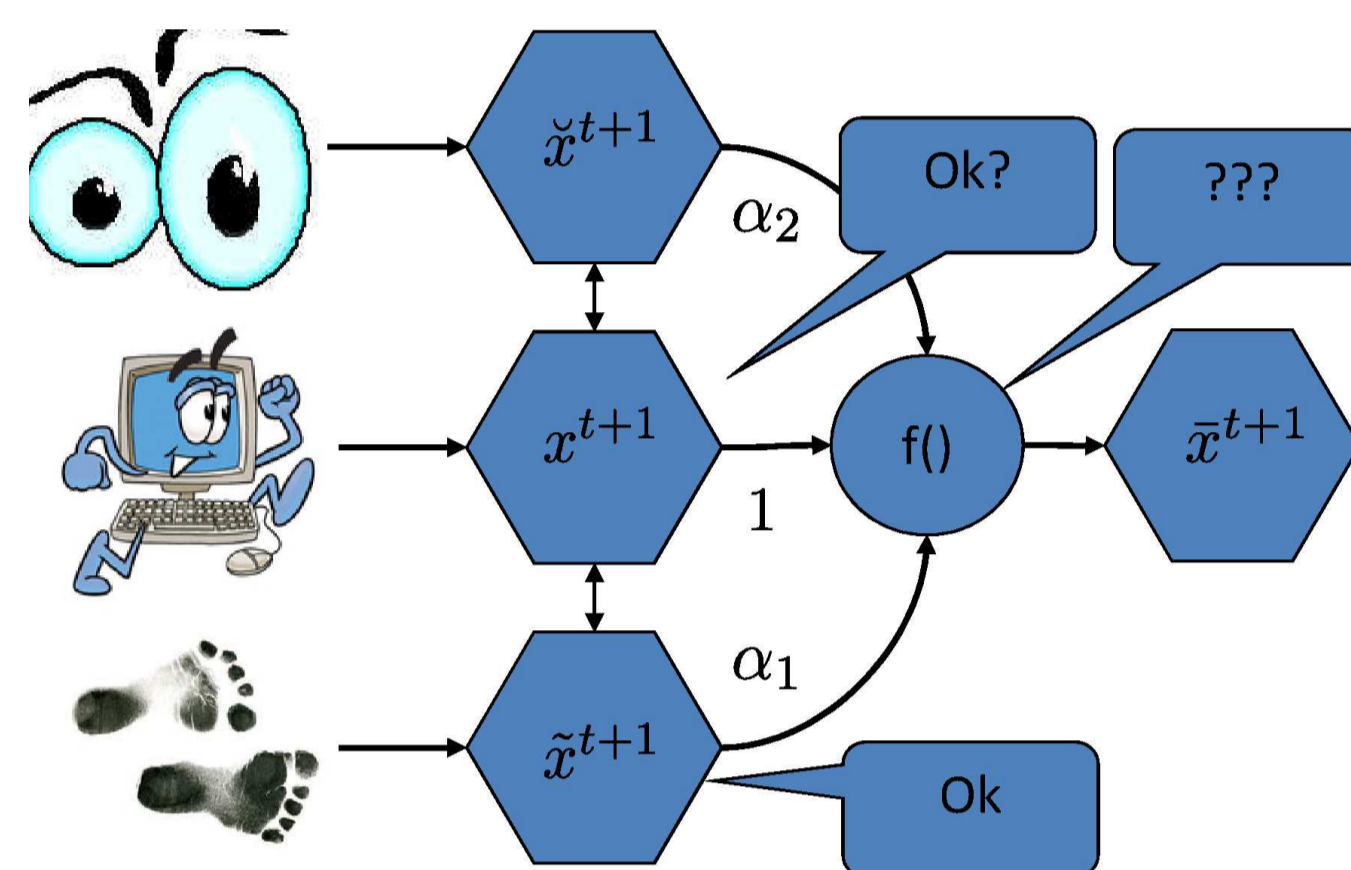

Figure 2: *Update of the internal model. The model's own prediction, sensory modalities and the motor signals contribute to the appropriate update according to their time-dependent reliability.*
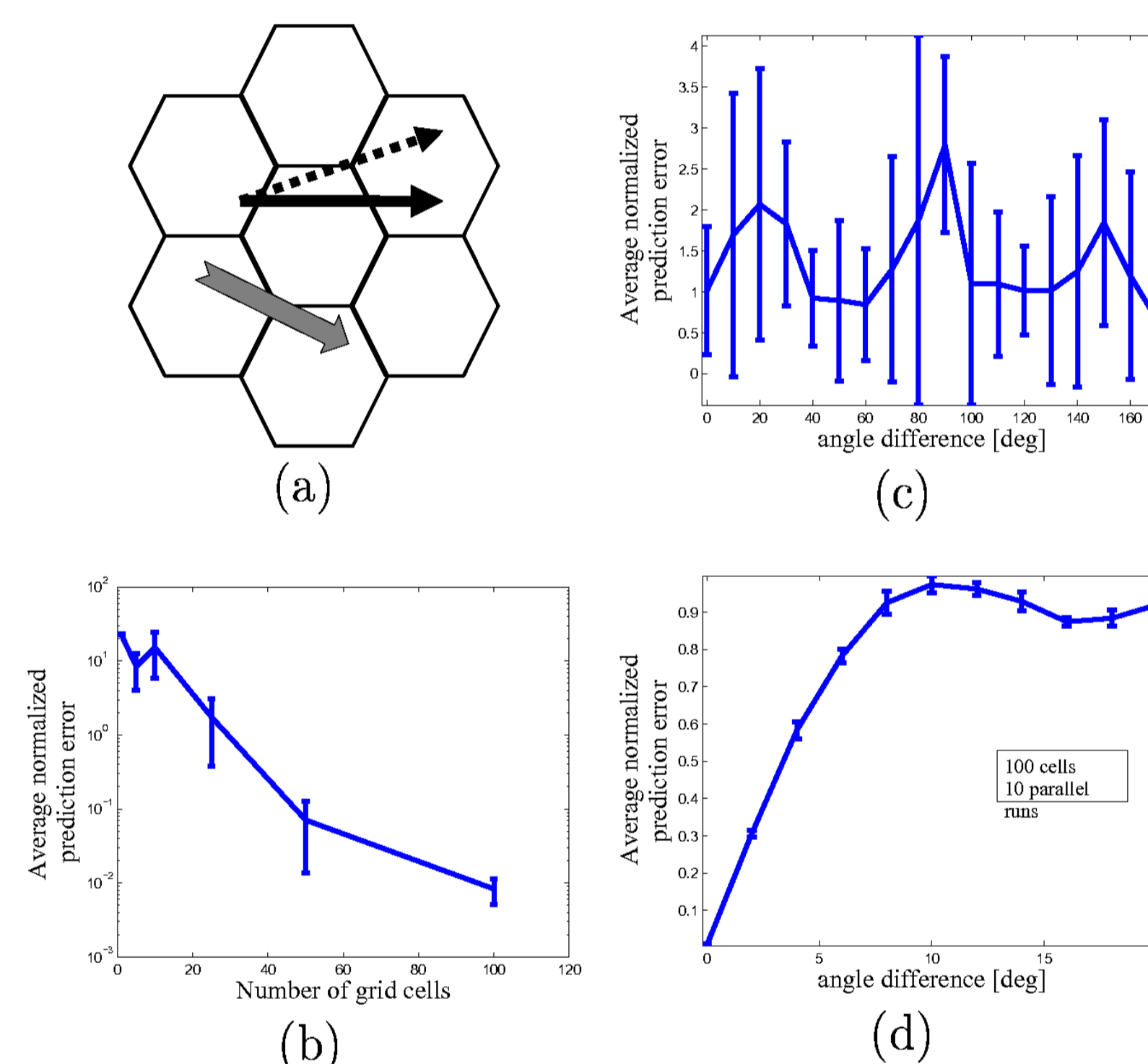

(a)

(b)

(c)

(d)

Figure 3: *Prediction of grid activity based on previous activity of units with similar direction preference. (a) the cartoon displays the training (dark, solid arrow) and the testing (dashed arrow) directions, while the thick notched arrow shows the preferred direction of the grid unit whose spatial activity pattern is denoted by the hexagon. During training only random, consecutive point pairs were used (as an AR(1) process is assumed), but during testing, the system had to predict the activity along the whole trajectory. (b) Average prediction error as a number of grid cells in the set. (c) Prediction error vs angle difference between training and testing tracks. (N=30). (d) Prediction error vs angle difference between the preferred and the training(testing) directions. (N=100)*

In most models on sensory motor maps or multimodal sensory fusion, this integration is thought to be realized by multiplicative *gain field control* mechanisms [12]. However, if neither of the modalities can be full trusted, these mechanisms may fail. We suggest instead a common low-dimensional representation onto which all modalities can be simultaneously projected to update the state of the internal dynamical system (Fig 2). In this formalism the outcome of the internal model can also be seen as a "sensory" modality: $\bar{x}^+ = f\{\alpha_1[Fx + \theta(|\hat{x}|)F_1(\hat{x} - x)] + \alpha_2[Fx + \theta(|\tilde{x}|)F_2(\tilde{x} - x)] + \ldots\}$. $\theta()$ denotes Heaviside step function to switch on and off the different terms of the r.h.s according to the presence of their argument. If no modalities can be accessed then the best guess is the prediction of the internal model ($\bar{x}^+ = f\{Fx\}$). The reliability of the terms are introduced through the values of α_i ($\sum_i \alpha_i = 1$). Differences between the modalities and the internal model's predictions are transmitted by matrices F_i (the simplest scenario, $Fx \approx \sum_i \alpha_i F^i x$) and $f()$ is a nonlinear transformation to keep x bounded. As a first step toward the design of this controlled dynamical system with probabilistic switching, first we need to show if periodic grid activity can be

learnt to extrapolate in time. In the simulations we used artificially generated (perfectly ordered) grids (multi-peaked spatial distribution with peaks on a regular hexagrid) and sampled the activity along parallel tracks covering the receptive field. A further simplification was that grid activity was approximated by a well defined group of grid cells (10-100 units of similar direction selectivity) with the same orientation and spacing but differing offsets. At last, Hebbian learning is not yet imposed on the system, we simply wanted to see, if the task can be solved at all within the proposed framework. We implemented a Support Vector Machine algorithm for regression (SVR) [13] as one of the most efficient (in terms of training time and robustness to perturbations or noise) algorithm for predicting time series. During training subsequent activity patterns (sample pairs along trajectory) were used. During testing the system was driven by its own output (n step look ahead prediction). Error is then defined as the 2-norm (spectral norm) of the difference between the real and the predicted population activities normalized by the number of grid units and averaged over 10 parallel runs. Fig 3 shows how the error depend on a, the number of units in the group b, the angle difference between training (same as the preferred direction) and testing tracks and c, the angle difference between the preferred and the training direction (same as the testing direction). The results show that a modest number of units is enough to predict the output of one unit and activity can indeed be predicted in the 2D space.

3. Work in progress

Treating modalities as control over the internal representation may lead to a better understanding of vestibular and motor feedback onto the HF. The rich recurrent collateral system of CA3 has not yet included in the model, but it is also assumed to effect grid formation. At last, it would be important to explain the role of grid representation in encoding long term memories in the neocortex.

References

[1] Lőrincz, A., Szirtes, G.: Here and now: how time segments may become events in the hippocampus. Neural Networks **22** (2009) 738–747

[2] Póczos, B., Lőrincz, A.: Non-combinatorial estimation of independent autoregressive sources. Neurocomputing **69** (2006) 2416–2419

[3] Jutten, C., Herault, J.: Blind separation of sources. Part I: An adaptive algorithm based on neuromimetic architecture. Signal Processing **24** (1991) 1–10

[4] Fyhn, M., Molden, S., Witter, M., Moser, E., Moser, M.B.: Spatial representation in the entorhinal cortex. Science **305** (2004) 1258–1264

[5] McNaughton, B.L., Battaglia, F.P., Jensen, O., Moser, E.I., Moser, M.B.: Path integration and the neural basis of the 'cognitive map'. Nature Reviews Neuroscience **7** (2006) 663–678

[6] Barry, C., Hayman, R., Burgess, N., Jeffery, K.J.: Experience-dependent rescaling of entorhinal grids. Nature Neuroscience **10** (2007) 682–684

[7] Kubie, J.L., Fenton, A.A.: Heading-vector navigation based on head-direction cells and path integration. Hippocampus **19** (2009) 456–479

[8] Lavenex, P., Lavenex, P.B., Amaral, D..: Spatial relational learning persists following neonatal hippocampal lesions in macaque monkeys. Nature Neuroscience **10** (2007) 234–239

[9] Giocomo, L.M., Hasselmo, M.E.: Computation by oscillations: Implications of experimental data for theoretical models of grid cells. Hippocampus **18** (2008) 1186–1199

[10] Burgess, N., Barry, C., O'Keefe, J.: An oscillatory interference model of grid cell firing. Hippocampus **17** (2007) 801–812

[11] Jeffery, K.J.: Integration of the sensory inputs to place cells: What, where, why, and how? Hippocampus **17** (2007) 775–785

[12] Salinas, E., J.Sejnowski, T.: Gain modulation in the central nervous system: where behavior, neurophysiology, and computation meet. Neuroscientist **7** (2001) 430

[13] Drucker, H., Burges, C.J., Kaufman, L., Smola, A., Vapnik, V.: Support vector regression machines. In: Advances in Neural Information Processing Systems 9. NIPS 1996, MIT Press (1997) 155–161